REACT: The Riskmap Evaluation and Coordination Terminal

by

Abraham Quintero

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2019

Author . . . .
Department of Electrical Engineering and Computer Science
August 20, 2019

Certified by . . . .
Miho Mazereeuw
Associate Professor
Thesis Supervisor

Accepted by . . . .
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Theses

REACT: The Riskmap Evaluation and Coordination Terminal

by

Abraham Quintero

Submitted to the Department of Electrical Engineering and Computer Science
on August 20, 2019, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Computer Science and Engineering

Abstract

The United Nations Office for Disaster Risk Reduction (UNDRR) states that economic losses due to natural disasters have risen 151 percent in the past 20 years. Of these disasters, floods are the most common. The Sendai Framework for Disaster Risk Reduction was created by the UNDRR in order to chart goals for future risk mitigation; among its seven global targets is increasing the availability of disaster risk information and assessment systems. Disaster information systems use state of the art techniques such as remote sensing in order to mitigate damages from natural and man made hazards.

More developed countries utilize networks of advanced sensors and ahead of time mapping in order to facilitate emergency responses; however, such systems are not available in developing countries due to cost limitations. The widespread proliferation of smart phones and social media use in developing countries means that citizens can be used as sensors by reporting disaster information online. The Riskmap system was developed by the Urban Risk Lab at MIT in order to gather citizen report streams. Such citizen disaster reports have two issues: a large influx of reports can cause information overload in emergency operations centers, which makes it difficult to summarize the situation. Machine learning has previously been used in order to analyze and simplify information for human consumption. This work seeks to use novel machine learning techniques to fully utilize crowdsourced social media reports gathered using the Riskmap system.

First we establish the motivation for using citizens as sensors and analyzing this noisy data using machine learning. We then review different machine learning techniques that have been used in crisis information systems, including those that also utilize social media. Finally a novel ensemble learning model is presented that can accurately predict large flood events from crowdsourced data.

Thesis Supervisor: Miho Mazereeuw
Title: Associate Professor

Acknowledgments

To Aditya and Miho- not one page of this thesis could have been written without your help and your guidance. Thank you so much for your patience and your expertise.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Flooding is the most common natural disaster in the world [3]. Flood related deaths account for half of all deaths from natural disasters [5]. Although flooding impacts both developed and developing countries, developing nations face much worse consequences as a result of flooding since they lack resources to adequately mitigate disasters [7]. Unregulated urbanization, rising population and climate change all contrive to increase the rate at which floods occur in developing megacities; furthermore, there is little data about these disasters [2]. Data scarcity makes it hard to pinpoint where to direct aid during disasters and where to make infrastructure improvements after disasters [8].

Government and NGOs work together to mitigate damages from flooding. Citizens look for relevant flood information and try to reduce their risk. Information is at the core of this interaction; however, data scarcity makes it hard for emergency personnel to optimize their use of resources, while citizens have an abundance of information about their surroundings but must be careful not to trust incorrect or outdated information about broader areas [6]. The natural solution is for citizens on social media to submit real time reports to the Emergency Operations Center (EOC), which is tasked with using those reports to inform citizens. There is one problem with this solution, in times of crisis EOCs can suffer from information overload when they are presented with too much information.

The REACT system uses novel machine learning and human computer interaction research to reduce information overload in EOCs, thereby decreasing disaster response time. REACT learns how Emergency Operations Centers (EOCs) classify the severity of flood events given citizen submitted reports. REACT trains itself through a gamified simulation of a disaster event. During a real disaster, REACT digests social media reports and estimates how severely an event is impacting different areas of a city and thereby helps EOCs to respond in the best manner possible.

1.1 History of Disaster Informatics

Work in Mapping disasters Epidemiology John Snow's use of maps to find the source of Cholera outbreak in London[9].

For flooding: [1]

Technology can help disaster response; however, it also has the ability to cause information overload[10]

1.2 The Riskmap System

1.2.1 Motivation for crowdsourced data

Citizens as sensors Geosocial intelligence Holderness [4] 1.2.

Quarantelli emphaszied that 'management of hazards is fundamentally social in nature and not something that can be achieved strictly through technological upgrading' [10] yet social media brings human behavior into a machine readable format that can be used to provide further information during disasters.

1.2.2

Chapter 2

Previous Work

2.1 Machine Learning in Crisis Informatics

2.1.1 Passive Listening

Most of the work in this area has been done by passively listening to twitter posts or facebook comments.

Problems with that approach

2.1.2 On Image Data

2.1.3 On Text Data

CrisisNLP has a huge datasets

2.1.4 Ensemble Data Models

Chapter 3

Methodology

3.1 Data Description

The Riskmap system allows citizens to easily submit disaster reports 1.2; as such it has allowed the Urban Risk Lab at MIT to gather thousands of reports of real flooding in Indonesia and India.

3.1.1 Image Data

3.1.2 Text Data

3.1.3 Flood Height

3.1.4 Location Information

3.2

Appendix A

Tables

Table A.1: Armadillos

Armadillos	are
our	friends

Appendix B

Figures

Figure B-1: Armadillo slaying lawyer.

Figure B-2: Armadillo eradicating national debt.

Bibliography

[1] Mike Ahern, R. Sari Kovats, Paul Wilkinson, Roger Few, and Franziska Matthies. Global Health Impacts of Floods: Epidemiologic Evidence. *Epidemiologic Reviews*, 27(1):36–46, July 2005.

[2] F. K. S. Chan, C. Joon Chuah, A. D. Ziegler, M. Dąbrowski, and O. Varis. Towards resilient flood risk management for Asian coastal cities: Lessons learned from Hong Kong and Singapore. *Journal of Cleaner Production*, 187:576–589, June 2018.

[3] Faith Ka Shun Chan, Gordon Mitchell, Olalekan Adekola, and Adrian McDonald. Flood Risk in Asia's Urban Mega-deltas: Drivers, Impacts and Response. *Environment and Urbanization ASIA*, 3(1):41–61, March 2012.

[4] Tomas Holderness and Etienne Turpin. From Social Media to GeoSocial Intelligence: Crowdsourcing Civic Co-management for Flood Response in Jakarta, Indonesia. In Surya Nepal, Cécile Paris, and Dimitrios Georgakopoulos, editors, *Social Media for Government Services*, pages 115–133. Springer International Publishing, Cham, 2015.

[5] C. A Ohl. Flooding and human health. *BMJ*, 321(7270):1167–1168, November 2000.

[6] E. L. Quarantelli. Problematical aspects of the information/ communication revolution for disaster planning and research: Ten non-technical issues and questions. *Disaster Prevention and Management; Bradford*, 6(2):94–106, 1997.

[7] E. L. Quarantelli. Urban Vulnerability to Disasters in Developing Countries: Managing Risks. In Alcira Kreimer, Margaret Arnold, and Anne Carlin, editors, *Building Safer Cities: The Future of Disaster Risk*, pages 211–231. Disaster Risk Management Series., 2003.

[8] Irfan Ahmad Rana and Jayant K. Routray. Multidimensional Model for Vulnerability Assessment of Urban Flooding: An Empirical Study in Pakistan. *International Journal of Disaster Risk Science; Heidelberg*, 9(3):359–375, September 2018.

[9] Simon Rogers. John Snow's data journalism: The cholera map that changed the world. *The Guardian*, March 2013.

[10] Kathleen J. Tierney, Michael K. Lindell, and Ronald W. Perry. *Facing the Unexpected: Disaster Preparedness and Response in the United States*. Joseph Henry Press, November 2001.